A transcription regulatory network within the ACE2 locus may promote a pro-viral environment for SARS-CoV-2 by modulating expression of host factors.

Tayaza Fadason[1,2¶], Sreemol Gokuladhas[1¶], Evgeniia Golovina[1¶], Daniel Ho[1], Sophie Farrow[1], Denis Nyaga[1], Hong Pan[3], Neerja Karnani[3,4], Conroy Wong[5], Antony Cooper[6,7], William Schierding[1,2], Justin M. O'Sullivan[1,2,8ǂ].

[1] The Liggins Institute, The University of Auckland, Auckland, New Zealand
[2] Maurice Wilkins Centre for Molecular Biodiscovery, Auckland, New Zealand
[3] Singapore Institute for Clinical Sciences (SICS), Agency for Science, Technology and Research (A*STAR), Singapore, Singapore
[4] Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore (NUS), Singapore
[5] Respiratory Medicine, Middlemore Hospital and University of Auckland, Auckland, New Zealand
[6] Australian Parkinsons Mission, Garvan Institute of Medical Research, Sydney, New South Wales, Australia
[7] St Vincent's Clinical School, UNSW Sydney, Sydney, New South Wales, Australia
[8] Brain Research New Zealand, The University of Auckland, Auckland, New Zealand.
[¶] Equal contributors
[ǂ] Corresponding author: Justin M. O'Sullivan, The Liggins Institute, The University of Auckland, Phone:+64 9 9239868, justin.osullivan@auckland.ac.nz

Abstract

Introduction: A novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was recently identified as the pathogen responsible for the COVID-19 outbreak. SARS-CoV-2 triggers severe pneumonia, which leads to acute respiratory distress syndrome and death in severe cases. As reported, SARS-CoV-2 is 80% genetically identical to the 2003 SARS-CoV virus. Angiotensin-converting enzyme 2 (ACE2) has been identified as the main receptor for entry of both SARS-CoV and SARS-CoV-2 into human cells. ACE2 is normally expressed in cardiovascular and lung type II alveolar epithelial cells, where it positively modulates the RAS system that regulates blood flow, pressure, and fluid homeostasis. Thus, virus-induced reduction of ACE2 gene expression is considered to make a significant contribution to severe acute respiratory failure. Chromatin remodeling plays a significant role in the regulation of ACE2 gene expression and the activity of regulatory elements within the genome.

Methods: Here, we integrated data on physical chromatin interactions within the genome organization (captured by Hi-C) with tissue-specific gene expression data to identify spatial expression quantitative trait loci (eQTLs) and thus regulatory elements located within the ACE2 gene.

Results: We identified regulatory elements within *ACE2* that control the expression of PIR, CA5B, and VPS13C in the lung. The gene products of these genes are involved in inflammatory responses, *de novo* pyrimidine and polyamine synthesis, and the endoplasmic reticulum, respectively.

Conclusion: Our study, although limited by the fact that the identification of the regulatory interactions is putative until proven by targeted experiments, supports the hypothesis that viral silencing of *ACE2* alters the activity of gene regulatory regions and promotes an intra-cellular environment suitable for viral replication.

Keywords: COVID-19, SARS-CoV-2, ARDS, ACE2

Introduction

Within months of the first reports [1], the COVID-19 outbreak has become a pandemic infecting and killing thousands of people worldwide [2]. COVID-19 is an infectious disease associated with acute respiratory distress syndrome (ARDS) that is caused by SARS-CoV-2, a Betacoronavirus that is 80% identical to the SARS-CoV virus [3]. Betacoronaviruses, including SARS-CoV, Murine Hepatic Virus (MHV), and SARS-CoV-2, utilize the ACE2 protein for cell entry [4,5]. The Spike protein on SARS-CoV-2 has a 10 to 20 fold higher affinity for the ACE2 protein than its SARS-CoV homologue [3,6].

The ACE2 protein is highly expressed in cardiovascular and lung type II alveolar epithelial cells [3,7,8], where ACE2 is a primary modulator of the renin–angiotensin (RAS) system that regulates blood flow, pressure and fluid homeostasis [9]. The ACE2 protein and the products of the reactions it catalyzes have also been implicated in immune responses and anti-inflammatory pathways [10–12].

SARS-CoV infection reduces *ACE2* gene expression and this is thought to contribute to severe acute respiratory failure [4] by triggering an imbalance in the RAS system that causes a loss of fluid homeostasis, induces inflammatory responses [10,13,14], and results in severe acute injury in heart and lung [3,15,16]. As mentioned above, both SARS-CoV and SARS-CoV-2 utilize the ACE2 protein for cell entry. Poor prognoses in elderly SARS-CoV-2 patients (≥65 years old) are frequently associated with a pre-existing reduction in *ACE2* expression and imbalance in ACE2-related host derived pathways [17,18]. *ACE2* is an X-linked gene whose expression is regulated by chromatin structure. Brg1, a chromatin remodeler, and the FoxM1 transcription factor recognize the *ACE2* promoter and reduce expression through a mechanism involving structural chromatin changes [19]. This control is complex, as illustrated by the finding that *ACE2* gene escapes X chromosome-inactivation and shows a heterogeneous sex bias that is tissue dependent [20].

Chromatin structure in the nucleus involves non-random folding of DNA on different scales [21]. This folding and the resulting contacts that form are dynamic, and can be disrupted (*e.g.* by genetic variation) leading to altered enhancer-promoter interactions that result in changes in gene expression [22]. Changing chromatin structure rewires interactions between regulatory elements and the genes they control. Theoretically and practically, each component of this change contributes to the observed pathogenesis [23,24], and can lead to developmental disorders[25] and cancer [26–28].

Virus-induced chromatin changes at the ACE2 locus could induce expression changes in additional genes regulated by elements located within this locus and thus may alter/modulate host factors important for SARS-CoV-2 replication. How can you identify the elements within a gene regulatory network like this? One approach to identify the networks that form between regulatory elements and the genes they control is to use information on the physical interactions that are captured occurring between the elements. Physical interactions between two sites can be captured and identified using Hi-C [29,30]. We have used this insight to develop a discovery-based pipeline (CoDeS3D; S1 Fig) [23]. Our approach uses genetic variation (*e.g.* single nucleotide polymorphisms) to identify changes in gene expression and thus determine if a region that physically contacts a gene contains a regulatory element. This enables the rapid identification of the regulatory networks that form in cells and tissues (*e.g.* [23,31]).

We hypothesized that *ACE2* and its flanking region contained regulatory elements that coordinate the expression of other genes, and that virus induced chromatin changes at ACE2 inadvertently modulate host factors that promote viral replication. Here we undertook an indepth characterization of the regulatory control regions within *ACE2* and their activity in lung tissue. Regulatory elements located within *ACE2* affect the expression levels of the *PIR* and *CA5B* genes. PIR and CA5B are involved in NF-kB regulation and pyrimidine synthesis, respectively. *VPS13C,* encoding a factor required for late stage endosome maturation, is also controlled by a putative enhancer located in intron 11 of *BMX*, adjacent to *ACE2*. We propose that *ACE2* repression by SARS-CoV-2 trips a chromatin-based switch that coordinates the activity of these regulatory elements and thus the genes they control. Collectively, these changes inadvertently lead to the development of a pro-viral replication environment.

Methods

Identification of SNPs in the *ACE* locus

We selected all common single nucleotide polymorphisms (SNPs) from dbSNP (build153) with a minor allele frequency (MAF) > 1% that were located within chrX:15,519,996-15,643,106, which included the *ACE2* gene and its flanking region (hereafter *ACE2* locus). SNP positions are as reported for the human genome build hg38 release 75 (GRCh38).

Identification of tissue-specific SNP-gene spatial relationships in the *ACE* locus

We used the CoDeS3D algorithm [23] to identify putative spatial regulatory interactions for all SNPs at the *ACE2* locus (S1 Fig). CoDeS3D integrates data on physical chromatin interactions within the genome organization (captured by Hi-C) with tissue-specific gene expression data to identify spatial expression quantitative trait loci (eQTLs). To get lung-specific spatial connections, we identified SNP-gene pairs across lung-specific Hi-C libraries using published data for IMR90, A549, and NCI-H460 cell lines and lung tissue (GEO accession numbers GSE35156, GSE43070, GSE63525, GSE105600, GSE105725, GSE92819, GSE87112, S1 Table). We then queried GTEx for eQTL associations with lung tissue (dbGaP Accession phs000424.v8.p2, https://gtexportal.org/home/). The age of the GTEx lung sample donors peaks between 50-60 years (S2 Fig). SNPs were assigned to the appropriate Hi-C restriction fragments by digital digestion of the hg38 reference genome (matching the restriction enzyme from the Hi-C libraries: *Mbo*I or *Hin*dIII). All SNP-fragments were queried against the Hi-C databases to identify the distal DNA fragments with which they physically interact. For each distal fragment, which overlapped a gene coding region, a SNP-gene spatial connection was confirmed. There was no binning or padding around restriction fragments to obtain gene overlap. Spatial tissue-specific SNP-gene pairs with significant eQTLs (both cis-acting [<1Mb between the SNP and gene] and trans-acting eQTLs [>1Mb between the SNP and gene or on different chromosomes]; FDR adjusted $p < 0.05$) within the lung were subsequently identified by querying GTEx v8 lung tissue (UBERON:0008952).

URLs

GEO database: https://www.ncbi.nlm.nih.gov/geo/

CoDeS3D pipeline: https://github.com/Genome3d/codes3d-v2

GTEx Portal: https://gtexportal.org/home/

GUSTO study: http://www.gusto.sg/

Data and code availability

All python and R scripts used for data analysis and visualization are available at https://github.com/Genome3d/ACE2-regulatory-network. R version 3.5.2 and RStudio version 1.2.5033 were used for all R scripts. All python scripts used Python 3.7.6.

Results

The *ACE2* locus harbors regulatory variants that control SARS-CoV-2 relevant cellular functions:

We tested 367 common SNPs located across the *ACE2* locus (chrX: 15,519,996-15,643,106) for their potential to act as spatial eQTLs. None of the common SNPs we tested affected *ACE2* expression levels in lung tissue (S2 Table).

The wider *ACE2* locus (chrX: 15,519,996-15,643,106; GRCh38/hg38) sits within a topologically associating domain (TAD) that is conserved across some tissues, *e.g.* IMR90 (Fig 1A). Therefore, it was not surprising that we identified control elements within this *ACE2* locus (Fig 1A). The distribution of targets for the putative control elements we identified is consistent with previous studies that show that while the majority of significant eQTLs fall within 100 kb of the transcription start site of a gene, only 60% of all eQTLs are upstream of the gene they regulate[32]. Notable amongst the elements we identified are long distance *trans*-regulatory interactions involving: 1) rs1399200:*VPS13C* (chr15:61,852,389-62,060,473; encodes vacuolar protein sorting-associated protein 13C); and 2) rs6632680:*PHKA2* (chrX:18,892,300-18,984,598; encodes phosphorylase kinase regulatory subunit alpha 2) (S2 Table).

We identified eighty genetic variants within the ACE2 locus as *cis*-acting spatial eQTLs that physically modulate the expression of genes *PIR* (encodes Pirin), *CA5BP1* (a pseudogene of CA5B), and *CA5B* (encodes mitochondrial carbonic anhydrase) in lung tissues (S2 Table). Fifty-eight SNPs located across the region are associated with increased expression of *PIR* ($\log_2$[aFC, allelic fold change] 0.462 ± 0.07) consistent with the elements they mark repressing PIR transcription. Eighteen SNPs located across the region are associated with decreased expression of *CA5B* ($\log_2$[aFC] -0.257 ± 0.005) consistent with the elements they mark enhancing *CA5B* transcription. These variants occurred in two clusters: 1) within the *ACE2* gene; and 2) within the *CLTRN* (*TMEM27*) gene – a known homologue of *ACE2*. Expression of *CA5BP1*, a pseudogene of *CA5B*, was also repressed ($\log_2$[aFC] -0.21 ± 0.01) by 6 SNPs within the *ACE2* locus. Within *ACE2* itself there were only control regions for the *PIR* and *CA5B* genes (Fig 1B).

The common variants that we tested show an unusual ancestry associated pattern of minor allele frequencies (Fig 1B). Specifically, the East Asian population (1K Genomes project) displays little variation across the bulk of the variants we analyzed. This observation is supported by measures of genetic diversity (F_{ST}) between the Indian, Chinese and Malay populations within the Growing Up in Singapore Towards healthy Outcomes (GUSTO) cohort (S3 Table). However, this pattern breaks down at several positions across the *ACE2* gene (including rs4646142, rs2285666, and rs2106809, which show significant selection towards the reference allele) in all of the tested populations, indicating potential selective pressure at these loci (Fig 1B). Notably, two of these variants alter potential transcription factor binding sites (*i.e.* rs2285666 alters HNF1, and Ncx motifs, rs2106809 alters a CEBPB motif; S4 Table). All three variants (rs4646142, rs2285666, and rs2106809) have previously been associated with allele, sex and ethnicity specific impacts on hypertension, blood pressure, hypertrophic cardiomyopathy, type 2 diabetes, myocardial infarction (reviewed in[33]). Moreover, the CEBPB motif is recognized by the CCAAT enhancer binding protein-β which has been implicated in inflammatory responses in lung carcinoma cells [34].

Discussion

We identified transcription regulatory elements for *CA5B* and *PIR* that are active in lung tissue and are located within the *ACE2* gene. We also identified a transcription regulatory element (located in the *BMX* gene, adjacent to *ACE2*) for the *PIR* and *VPS13C* genes. It is sterically impossible for a single DNA sequence to simultaneously be transcribed and regulate another gene through a physical connection. Therefore, we propose that SARS-CoV-2-induced chromatin-dependent repression of *ACE2* expression in lung enables the regulatory sites, repressing *PIR* and activating *CA5B*, to exhibit increased functionality in infected cells (Fig 2). We hypothesize that this regulatory change extends to coordinate changes in the expression of *VPS13C* and *PHKA2* in ways that promote viral proliferation. This host regulatory network has not evolved to benefit the virus but rather, these regulatory changes inadvertently produce an environment advantageous for the virus.

The *CA5B* gene encodes a mitochondrial carbonic anhydrase that catalyzes the reversible hydration of CO_2 in the lung. This reaction is important in mitochondria as it supplies HCO_3^- ions required by pyruvate carboxylase for gluconeogenesis, and by carbamoyl phosphate synthase 1 (CSP1) for pyrimidine biosynthesis (Fig 2B) [38–40]. Pyrimidines are important host factors critical for viral genomic replication, mRNA synthesis for protein translation, and phospholipid synthesis [41]. Inhibiting *de novo* pyrimidine biosynthesis impacts on SARS-CoV-2 replication [18]. CSP1 additionally produces a precursor for the biosynthesis of polyamines, small aliphatic molecules that play important roles in virus replication. Inhibition of polyamine biosynthesis significantly impaired replication of the Middle East Respiratory Syndrome [MERS] coronavirus [42]. Targeted inhibition of CA5B encoded carbonic anhydrase might therefore decrease levels of critical host factors pyrimidines and polyamines - critical host factors needed for SARS-CoV-2 replication. Intriguingly, similarities in the pathologies of SARS-CoV-2 infection and high altitude pulmonary edema (HAPE) have led to the suggestion that carbonic anhydrase inhibitors could be used to treat or prevent Covid-19 infection [43].

Interleukin expression is responsible for irreversible, pathological changes associated with SARS-CoV infection in the lung (*e.g.* [44]). Human coronavirus has been shown to fine-tune NF-κB signaling [45]. *PIR* encodes a non-heme iron binding protein that is a redox switch that modulates the binding of p65 (RelA) to NF-κB responsive promoters [46]. NF-κB regulates multiple immune function aspects, including the production of pro-inflammatory cytokines[47]. Therefore, it is notable that repressor regulatory sequences for *PIR* sit within the *ACE2* gene (Fig 2A). We postulate that the chromatin modifications that silence *ACE2* expression upon early stage infection activate the *PIR* repressor (Fig 2B). This reduces responsiveness of NF-κB, and thereby delays the expected and needed anti-viral response. Reduction in PIR expression would also reduce the impact of any changes to intra-cellular redox state caused by that SARS-CoV-2 infection, however little is known and future experiments are required to clarify this.

The enveloped Betacoronaviruses (MHV, SARS-CoV, SARS-CoV-2) gain entry to the cell through the endo/lysosomal pathway and require late endosomal maturation for fusion [48]. Therefore, it is interesting to speculate on the impact of coordinated changes to *VPS13C* expression. The VPS13 family are endoplasmic reticulum associated lipid transporters. VPS13C is proposed to act as a lipid transporter at organelle contact sites between (i) the endoplasmic reticulum (ER) and endolysosomes, and (ii) the ER and lipid droplets, where it transfers lipids, potentially bulk lipid transfer, between organelles to maintain lipid homeostasis and organelle functionality [49]. Increased *VPS13C* expression is predicted to increase the extent of contact and lipid transfer between these organelles. This in turn could enhance the virus's replication capacity and pathogenesis, as the ER plays both a physical and functional central role in the virus's capacity to replicate and form new viral progeny. Moreover, SARS-CoV extensively reorganizes the host cell's membranes infrastructure to produce a reticulo-vesicular network of modified ER to coordinate its replication cycle [50]. Alterations to ER-lipid droplet contacts mediated by VPS13C could support the virus's required expansion and re-organization of ER membranes by altering lipid flow through the ER [51]. Notably, cells infected with Hepatitis C virus (HCV), a positive-strand RNA virus like SARS-CoV-2, contain ER-derived membranous structures that contain significantly high levels of cholesterol, despite the ER of uninfected cells possessing relatively low cholesterol levels [52]. Increased VPS13C-mediated ER-endolysosomal contact sites could increase the capacity of endocytosed dietary cholesterol to be delivered to the ER and enhance the virus's ability to replicate. Pharmacological impairment of endolysosomal cholesterol efflux reduced HCV replication, [52] suggesting another possible therapeutic approach for investigation to slow SARS-CoV-2 replication. ER stress, impacted by changes to *VPS13C* expression, may also contribute to late infection stage NF-κB activation (reviewed in [53]).

The significance of the putative enhancer for *PHKA2*, which encodes the phosphorylase kinase regulatory subunit alpha 2, is unclear. Mutations in this gene have been linked to glycogen storage disorders and glucose metabolism. Thus, linkages can be drawn to the increased expression of *CA5B*, which impacts on gluconeogenesis. Notably, *PHKA2* was downregulated in plasma from individuals with hepatocellular-carcinoma caused by HCV infections [54]. Theoretically, chromatin remodeling in response to SARS-COV-2 infection could down-regulate *PHKA2* expression. However, there is a paucity of information linking this gene to viral infections or the lung and this conclusion requires additional experimental support.

SARS-CoV is known to repress *ACE2* expression [4]. *ACE2* regulation involves chromatin remodeling and structural chromatin changes [19]. Several of the regions that we identified overlapped or were adjacent to CTCF biding sites (*e.g.* rs1399200, which regulates *PIR* and *VPS13C*; rs6629111, which regulates *CA5B*; and sites [rs714205, rs1514280, rs4240157 and rs4646131] within *ACE2*[33]). We also note that the regulatory sites we identified included transcription factor binding sites for Ap-1, RXRA (a DNA-binding receptor involved in host-virus interactions) [55], GR or NR3C1 (a regulator of inflammation in asthma and COPD) [56], Pou2f2 (trans-activator of NR3C1)[57], and P300 (a chromatin modifier; S4 Table) [58]. Expression data within the search-based exploration of expression compendium (SEEK) supports a strong co-expression relationship between *ACE2* and *PIR* (lung cancer, ovarian tumor) and a weaker association with *CA5B* (ovarian cancer) (http://seek.princeton.edu/ [59]). However, the possible mechanism(s) that link ACE2 silencing to alterations associated with these regulatory regions remains unknown until empirically determined in lung cells in the presence/absence of real or simulated viral infection.

Whilst this study is novel and uses empirically derived data in the analyses, the observations are limited by the fact that the identification of the regulatory interactions is only putative until proven by targeted experiments. Ideally, the Hi-C datasets should be derived from matched tissues prior and post SARS-CoV-2 infection. Finally, the GTEx database has recognized limitations, including the ethnic diversity of the samples. These limitations will form the basis of future studies.

Conclusion

We identified putative regulatory regions in and surrounding *ACE2* that regulate the expression of *PIR*, *CA5B*, and *VPS13C* in the lung. We contend that viral induced chromatin-dependent repression of the *ACE2* gene increases the activity of these regulatory sites and promotes an intra-cellular environment suitable for viral replication. The altered gene products represent new targets for anti-SARS-CoV-2 therapeutics.

Acknowledgements

The authors would like to thank the Genomics and Systems Biology Group (Liggins Institute), Mark Hampton, and Elizabeth Ledgerwood for useful discussions. We would like to thank the funders of GTEx Project - common Fund of the Office of the Director of the National Institutes of Health, and by National Cancer Institute, National Human Genome Research Institute, National Heart, Lung, and Blood Institute, National Institute on Drug Abuse, National Institute of Mental Health, National Institute of Neurological Disorders and Stroke. The authors thank the GUSTO study group for comments on the manuscript. This work has been released as a pre-print.

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author contributions

TF, EG, SG and WS performed analyses and co-wrote the manuscript. DH, DN, SF, and AC performed literature searches and co-wrote the manuscript. HP and NK provide FST data from GUSTO and commented on the manuscript. CW reviewed the findings and commented on the manuscript. JOS led the study and co-wrote the manuscript.

Funding

SG and DH are recipients of scholarships funded by a Ministry of Business, Innovation and Employment Catalyst grant (New Zealand- Australia LifeCourse Collaboration on Genes, Environment, Nutrition and Obesity; UOAX1611) to JOS. JOS and WS are funded by a Royal Society of New Zealand Marsden Fund [Grant 16-UOO-072]. JOS and TF are funded by an HRC explorer grant (HRC19/774) to JOS. DN was supported by the Sir Colin Giltrap Liggins Institute Scholarship fund. AC received grant funding from the Australian government. GUSTO study is supported by Singapore National Research Foundation under its Translational and Clinical Research (TCR) Flagship Program administered by the Singapore Ministry of Health's National Medical Research Council (NMRC/TCR/004-NUS/2008; NMRC/TCR/012-NUHS/2014). SNP variant analysis in GUSTO cohort was supported by Industry Alignment Fund – Pre-positioning Programme (IAF-PP H17/01/a0/005, available to NK. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Chen J, Jiang Q, Xia X, Liu K, Yu Z, Tao W, et al. Individual Variation of the SARS-CoV2 Receptor ACE2 Gene Expression and Regulation. 2020.
2. WHO. Coronavirus disease 2019. World Heal Organ. 2020;2019: 2633. doi:10.1001/jama.2020.2633
3. Guo Y-R, Cao Q-D, Hong Z-S, Tan Y-Y, Chen S-D, Jin H-J, et al. The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak - an update on the status. Mil Med Res. 2020;7: 11. doi:10.1186/s40779-020-00240-0
4. Kuba K, Imai Y, Rao S, Gao H, Guo F, Guan B, et al. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus–induced lung injury. Nat Med. 2005;11: 875–879. doi:10.1038/nm1267
5. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell. 2020; 1–10. doi:10.1016/j.cell.2020.02.052
6. Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh C-L, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science. 2020;1263: 1260–1263. doi:10.1126/science.abb2507
7. Zhao Y, Zhao Z, Wang Y, Zhou Y, Ma Y, Zuo W. Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCov. bioRxiv. 2020; 2020.01.26.919985. doi:10.1101/2020.01.26.919985
8. Harmer D, Gilbert M, Borman R, Clark KL. Quantitative mRNA expression profiling of ACE 2, a novel homologue of angiotensin converting enzyme. FEBS Lett. 2002;532: 107–110. doi:10.1016/S0014-5793(02)03640-2
9. Tikellis C, Thomas MC. Angiotensin-converting enzyme 2 (ACE2) is a key modulator of the renin angiotensin system in health and disease. Int J Pept. 2012;2012. doi:10.1155/2012/256294
10. Rodrigues Prestes TR, Rocha NP, Miranda AS, Teixeira AL, Simoes-e-Silva AC. The Anti-Inflammatory Potential of ACE2/Angiotensin-(1-7)/Mas Receptor Axis: Evidence from Basic and Clinical Research. Curr Drug Targets. 2017;18: 1301–1313. doi:10.2174/1389450117666160727142401
11. Sodhi CP, Nguyen J, Yamaguchi Y, Werts AD, Lu P, Ladd MR, et al. A Dynamic Variation of Pulmonary ACE2 Is Required to Modulate Neutrophilic Inflammation in Response to Pseudomonas aeruginosa Lung Infection in Mice . J Immunol. 2019;203: 3000–3012. doi:10.4049/jimmunol.1900579
12. Steck AK, Dong F, Wong R, Fouts A, Liu E, Romanos J, et al. Improving prediction of type 1 diabetes by testing non-HLA genetic variants in addition to HLA markers. Pediatr Diabetes. 2014;15: 355–362. doi:10.1111/pedi.12092
13. Hashimoto T, Perlot T, Rehman A, Trichereau J, Ishiguro H, Paolino M, et al. ACE2 links amino acid malnutrition to microbial ecology and intestinal inflammation. Nature. 2012;487: 477–481. doi:10.1038/nature11228

14. Keidar S, Strizevsky A, Raz A, Gamliel-Lazarovich A. ACE2 activity is increased in monocyte-derived macrophages from prehypertensive subjects. Nephrol Dial Transplant. 2007;22: 597–601. doi:10.1093/ndt/gfl632

15. Imai Y, Kuba K, Rao S, Huan Y, Guo F, Guan B, et al. Angiotensin-converting enzyme 2 protects from severe acute lung failure. Nature. 2005;436: 112–116. doi:10.1038/nature03712

16. Imai Y, Kuba K, Penninger JM. The discovery of angiotensin-converting enzyme 2 and its role in acute lung injury in mice. Exp Physiol. 2008;93: 543–548. doi:10.1113/expphysiol.2007.040048

17. Xudong X, Junzhu C, Xingxiang W, Furong Z, Yanrong L. Age- and gender-related difference of ACE2 expression in rat lung. Life Sci. 2006;78: 2166–2171. doi:10.1016/j.lfs.2005.09.038

18. Xiong R, Zhang L, Li S, Sun Y, Ding M, Wang Y, et al. Novel and potent inhibitors targeting DHODH, a rate-limiting enzyme in de novo pyrimidine biosynthesis, are broad-spectrum antiviral against RNA viruses including newly emerged coronavirus SARS-CoV-2. bioRxiv. 2020; 2020.03.11.983056. doi:10.1101/2020.03.11.983056

19. Yang J, Feng X, Zhou Q, Cheng W, Shang C, Han P, et al. Pathological Ace2-to-Ace enzyme switch in the stressed heart is transcriptionally controlled by the endothelial Brg1–FoxM1 complex. Proc Natl Acad Sci. 2016;113: E5628–E5635. doi:10.1073/pnas.1525078113

20. Tukiainen T, Villani A-C, Yen A, Rivas MA, Marshall JL, Satija R, et al. Landscape of X chromosome inactivation across human tissues. Nature. 2017;550: 244–248. doi:10.1038/nature24265

21. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell. 2014;159: 1665–1680. doi:10.1016/j.cell.2014.11.021

22. Meier M, Grant J, Dowdle A, Thomas A, Gerton J, Collas P, et al. Cohesin facilitates zygotic genome activation in zebrafish. Development. 2018;145: dev156521. doi:10.1242/dev.156521

23. Fadason T, Schierding W, Lumley T, O’Sullivan JM. Chromatin interactions and expression quantitative trait loci reveal genetic drivers of multimorbidities. Nat Commun. 2018;9: 5198. doi:10.1038/s41467-018-07692-y

24. Liu X, Li YI, Pritchard JK. Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. Cell. 2019;177: 1022-1034.e6. doi:10.1016/j.cell.2019.04.014

25. Paulsen J, Liyakat Ali TM, Nekrasov M, Delbarre E, Baudement M-O, Kurscheid S, et al. Long-range interactions between topologically associating domains shape the four-dimensional genome during differentiation. Nat Genet. 2019;51: 835–843. doi:10.1038/s41588-019-0392-0

26. Rivas MA, Melnick AM. Role of chromosomal architecture in germinal center B cells

and lymphomagenesis. Current Opinion in Hematology. Lippincott Williams and Wilkins; 2019. pp. 294–302. doi:10.1097/MOH.0000000000000505

27. Yang M, Vesterlund M, Siavelis I, Moura-Castro LH, Castor A, Fioretos T, et al. Proteogenomics and Hi-C reveal transcriptional dysregulation in high hyperdiploid childhood acute lymphoblastic leukemia. Nat Commun. 2019;10: 1519. doi:10.1038/s41467-019-09469-3

28. Achinger-Kawecka J, Valdes-Mora F, Luu PL, Giles KA, Caldon CE, Qu W, et al. Epigenetic reprogramming at estrogen-receptor binding sites alters 3D chromatin landscape in endocrine-resistant breast cancer. Nat Commun. 2020;11: 320. doi:10.1038/s41467-019-14098-x

29. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science (80-). 2009/10/10. 2009;326: 289–293. doi:326/5950/289 [pii]10.1126/science.1181369

30. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014;159: 1665–1680. doi:10.1016/j.cell.2014.11.021

31. Gokuladhas S, Schierding W, Cameron-Smith D, Wake M, Scotter E, O'Sullivan JM. Shared regulatory pathways reveal novel genetic correlations between grip strength and neuromuscular disorders. Front Genet. 2020;In Press. doi:10.3389/fgene.2020.00393

32. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013;45: 580–585. doi:10.1038/ng.2653

33. Burrell LM, Harrap SB, Velkoska E, Patel SK. The ACE2 gene: its potential as a functional candidate for cardiovascular disease. Clin Sci. 2013;124: 65–76. doi:10.1042/CS20120269

34. Armstrong DA, Phelps LN, Vincenti MP. CCAAT enhancer binding protein-β regulates matrix metalloproteinase-1 expression in interleukin-1β-stimulated A549 lung carcinoma cells. Mol Cancer Res. 2009;7: 1517–1524. doi:10.1158/1541-7786.MCR-09-0082

35. Wang Y, Song F, Zhang B, Zhang L, Xu J, Kuang D, et al. The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. Genome Biol. 2018;19: 151. doi:10.1186/s13059-018-1519-9

36. Hinrichs a S, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, et al. The UCSC Genome Browser Database: update 2006. Nucleic Acids Res. 2006;34: D590-8. doi:10.1093/nar/gkj144

37. Ward LD, Kellis M. HaploReg v4: Systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. Nucleic Acids Res. 2016;44: D877–D881. doi:10.1093/nar/gkv1340

38. Ghandour MS, Parkkila AK, Parkkila S, Waheed A, Sly WS. Mitochondrial carbonic anhydrase in the nervous system: Expression in neuronal and glial cells. J Neurochem.

2000;75: 2212–2220. doi:10.1046/j.1471-4159.2000.0752212.x

39. Adeniran C, Hamelberg D. Redox-Specific Allosteric Modulation of the Conformational Dynamics of κb DNA by Pirin in the NF-κB Supramolecular Complex. Biochemistry. 2017;56: 5002–5010. doi:10.1021/acs.biochem.7b00528

40. Kim J, Hu Z, Cai L, Li K, Choi E, Faubert B, et al. CPS1 maintains pyrimidine pools and DNA synthesis in KRAS/LKB1-mutant lung cancer cells. Nature. 2017;546: 168–172. doi:10.1038/nature22359

41. Löffler M, Fairbanks LD, Zameitat E, Marinaki AM, Simmonds HA. Pyrimidine pathways in health and disease. Trends Mol Med. 2005;11: 430–7. doi:10.1016/j.molmed.2005.07.003

42. Mounce BC, Cesaro T, Moratorio G, Hooikaas PJ, Yakovleva A, Werneke SW, et al. Inhibition of Polyamine Biosynthesis Is a Broad-Spectrum Strategy against RNA Viruses. Pfeiffer J, editor. J Virol. 2016;90: 9683–9692. doi:10.1128/JVI.01347-16

43. Solaimanzadeh I. Acetazolamide, Nifedipine and Phosphodiesterase Inhibitors: Rationale for Their Utilization as Adjunctive Countermeasures in the Treatment of Coronavirus Disease 2019 (COVID-19). Cureus. 2020;2019. doi:10.7759/cureus.7343

44. Yoshikawa T, Hill T, Li K, Peters CJ, Tseng C-TK. Severe Acute Respiratory Syndrome (SARS) Coronavirus-Induced Lung Epithelial Cytokines Exacerbate SARS Pathogenesis by Modulating Intrinsic Functions of Monocyte-Derived Macrophages and Dendritic Cells. J Virol. 2009;83: 3039–3048. doi:10.1128/jvi.01792-08

45. Poppe M, Wittig S, Jurida L, Bartkuhn M, Wilhelm J, Müller H, et al. The NF-κB-dependent and -independent transcriptome and chromatin landscapes of human coronavirus 229E-infected cells. Enjuanes L, editor. PLOS Pathog. 2017;13: e1006286. doi:10.1371/journal.ppat.1006286

46. Liu F, Rehmani I, Esaki S, Fu R, Chen L, de Serrano V, et al. Pirin is an iron-dependent redox regulator of NF- B. Proc Natl Acad Sci. 2013;110: 9722–9727. doi:10.1073/pnas.1221743110

47. Liu T, Zhang L, Joo D, Sun S-C. NF-κB signaling in inflammation. Signal Transduct Target Ther. 2017;2: 17023. doi:10.1038/sigtrans.2017.23

48. Burkard C, Verheije MH, Wicht O, van Kasteren SI, van Kuppeveld FJ, Haagmans BL, et al. Coronavirus Cell Entry Occurs through the Endo-/Lysosomal Pathway in a Proteolysis-Dependent Manner. Perlman S, editor. PLoS Pathog. 2014;10: e1004502. doi:10.1371/journal.ppat.1004502

49. Kumar N, Leonzino M, Hancock-Cerutti W, Horenkamp FA, Li P, Lees JA, et al. VPS13A and VPS13C are lipid transport proteins differentially localized at ER contact sites. J Cell Biol. 2018;217: 3625–3639. doi:10.1083/jcb.201807019

50. Knoops K, Kikkert M, Worm SHE van den, Zevenhoven-Dobbe JC, van der Meer Y, Koster AJ, et al. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. PLoS Biol. 2008;6: e226. doi:10.1371/journal.pbio.0060226

51. Lee J-Y, Cortese M, Haselmann U, Tabata K, Romero-Brey I, Funaya C, et al. Spatiotemporal Coupling of the Hepatitis C Virus Replication Cycle by Creating a Lipid Droplet- Proximal Membranous Replication Compartment. Cell Rep. 2019;27: 3602-3617.e5. doi:10.1016/j.celrep.2019.05.063

52. Stoeck IK, Lee J-Y, Tabata K, Romero-Brey I, Paul D, Schult P, et al. Hepatitis C Virus Replication Depends on Endosomal Cholesterol Homeostasis. J Virol. 2018;92. doi:10.1128/JVI.01196-17

53. Fung TS, Liu DX. Coronavirus infection, ER stress, apoptosis and innate immunity. Front Microbiol. 2014;5: 1–13. doi:10.3389/fmicb.2014.00296

54. Zekri ARN, Hafez MM, Bahnassy AA, Hassan ZK, Mansour T, Kamal MM, et al. Genetic profile of Egyptian hepatocellular-carcinoma associated with hepatitis C virus Genotype 4 by 15 K cDNA microarray: Preliminary study. BMC Res Notes. 2008;1. doi:10.1186/1756-0500-1-106

55. Raney AK, Johnson JL, Palmer CN, McLachlan A. Members of the nuclear receptor superfamily regulate transcription from the hepatitis B virus nucleocapsid promoter. J Virol. 1997;71: 1058–1071. doi:10.1128/JVI.71.2.1058-1071.1997

56. Barnes PJ. Corticosteroid Resistance in Airway Disease. Proc Am Thorac Soc. 2004;1: 264–268. doi:10.1513/pats.200402-014MS

57. Préfontaine GG, Walther R, Giffin W, Lemieux ME, Pope L, Haché RJG. Selective Binding of Steroid Hormone Receptors to Octamer Transcription Factors Determines Transcriptional Synergism at the Mouse Mammary Tumor Virus Promoter. J Biol Chem. 1999;274: 26713–26719. doi:10.1074/jbc.274.38.26713

58. Yoneyama M. Direct triggering of the type I interferon system by virus infection: activation of a transcription factor complex containing IRF-3 and CBP/p300. EMBO J. 1998;17: 1087–1095. doi:10.1093/emboj/17.4.1087

59. Zhu Q, Wong AK, Krishnan A, Aure MR, Tadych A, Zhang R, et al. Targeted exploration and analysis of large cross-platform human transcriptomic compendia. Nat Methods. 2015;12: 211–214. doi:10.1038/nmeth.3249

Figure legends

Fig 1. Elements located within and surrounding the *ACE2* locus regulate the lung-specific expression of *PIR*, *CA5B*, *CA5BP1*, *VPS13C*, and *PHKA2*. (A) Common genetic variants (SNPs) located within the *ACE2* locus form spatial *cis*-acting regulatory interactions with *PIR*, *CA5BP1*, and *CA5B* across sub-TAD boundaries on chrX:15,300,000-15,600,000. Inter-TAD *trans*-acting interactions regulate *PHKA2* (3.2 Mb away) and *VPS13C* (located on chromosome 15). Visualization of TAD and chromatin interactions was performed using the 3D genome browser (http://yuelab.org/)[35] and UCSC browser's interact tool (http://genome.ucsc.edu)[36], respectively. (B) Within *ACE2*, MAFs for the SNPs that tag the regulatory sites showed significant bias in four different populations (*i.e.* African [AFR], Ad Mixed American [AMR], East Asian [ASN] and European [EUR]) at one *PIR* (rs714205) and three *CA5B* regulatory sites (rs4646142, rs2285666, and rs2106809), consistent with selection acting on these loci. MAFs were obtained from HaploReg v4.1 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) [37].

Fig 2. Hypothesis: SARS-CoV-2 infection is associated with an *ACE2* dependent switch that alters expression of proteins that promote an environment for viral proliferation in the lung. (A) In middle-aged non-infected individuals, control regions within *ACE2* are capable of downregulating the expression of *PIR*, which is involved in the NF-κB pathway. Enhancer elements within *ACE2* are poised to upregulate *CA5B* expression, which encodes an enzyme important for pyrimidine synthesis. In addition to this, an enhancer region within the *BMX* gene (still within the same TAD) contributes to *VPS13C* regulation. (B) We hypothesize that upon viral infection, SARS-COV-2 represses *ACE2* expression, which increases the activity of the *PIR* repressor and *CA5B* enhancer. This results in a reduction in the production of PIR - the redox switch necessary for NF-κB activation, while also increasing pyrimidine synthesis, which is necessary for viral replication.

Supplementary Tables

S1 Table. Lung-specific Hi-C libraries used in the analysis

S2 Table. Lung-specific spatial SNP-gene relationships in the *ACE* locus

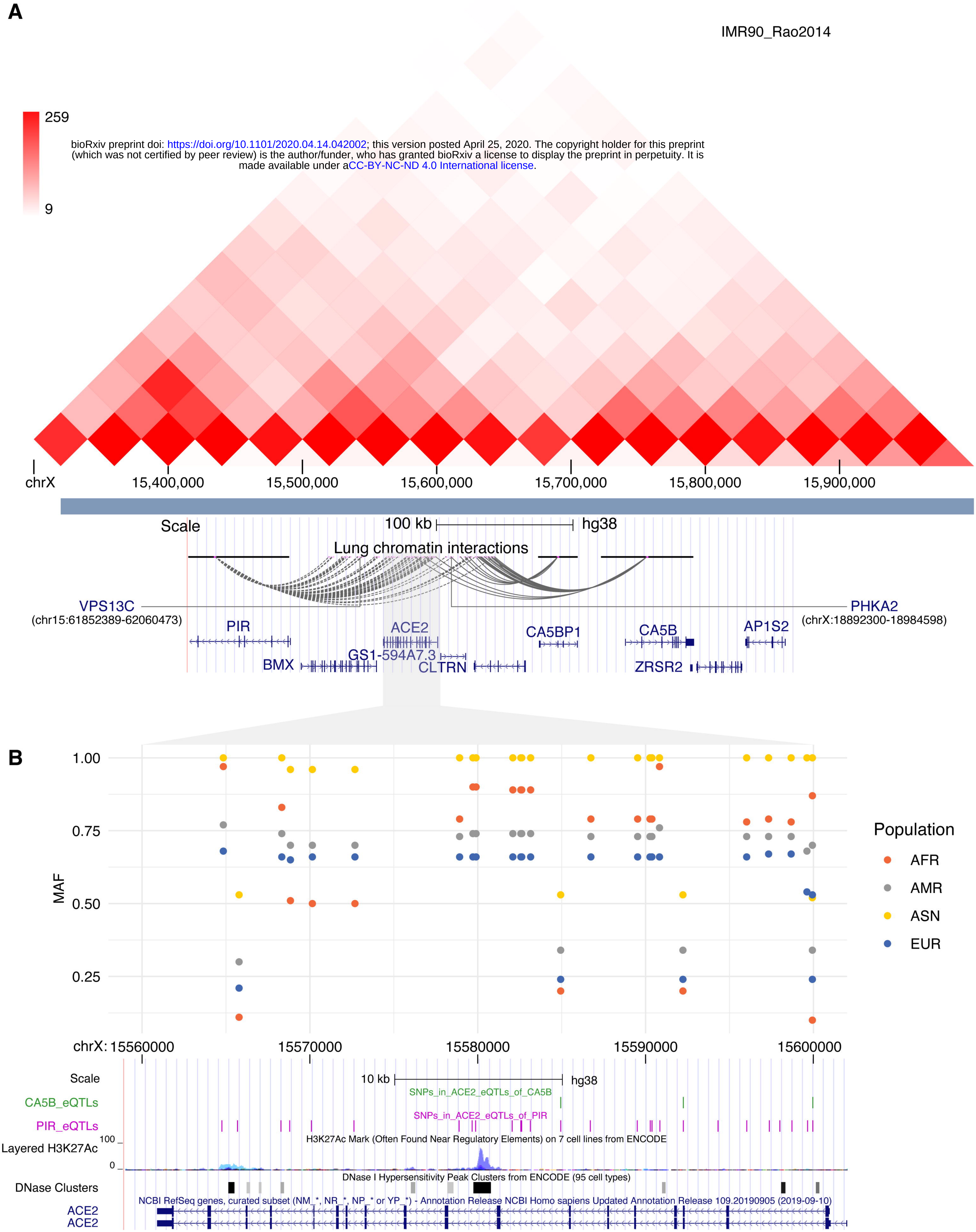
S3 Table. Genetic diversity estimate (Fst) across *ACE2* in the Indian, Malay and Chinese populations in the GUSTO cohort.

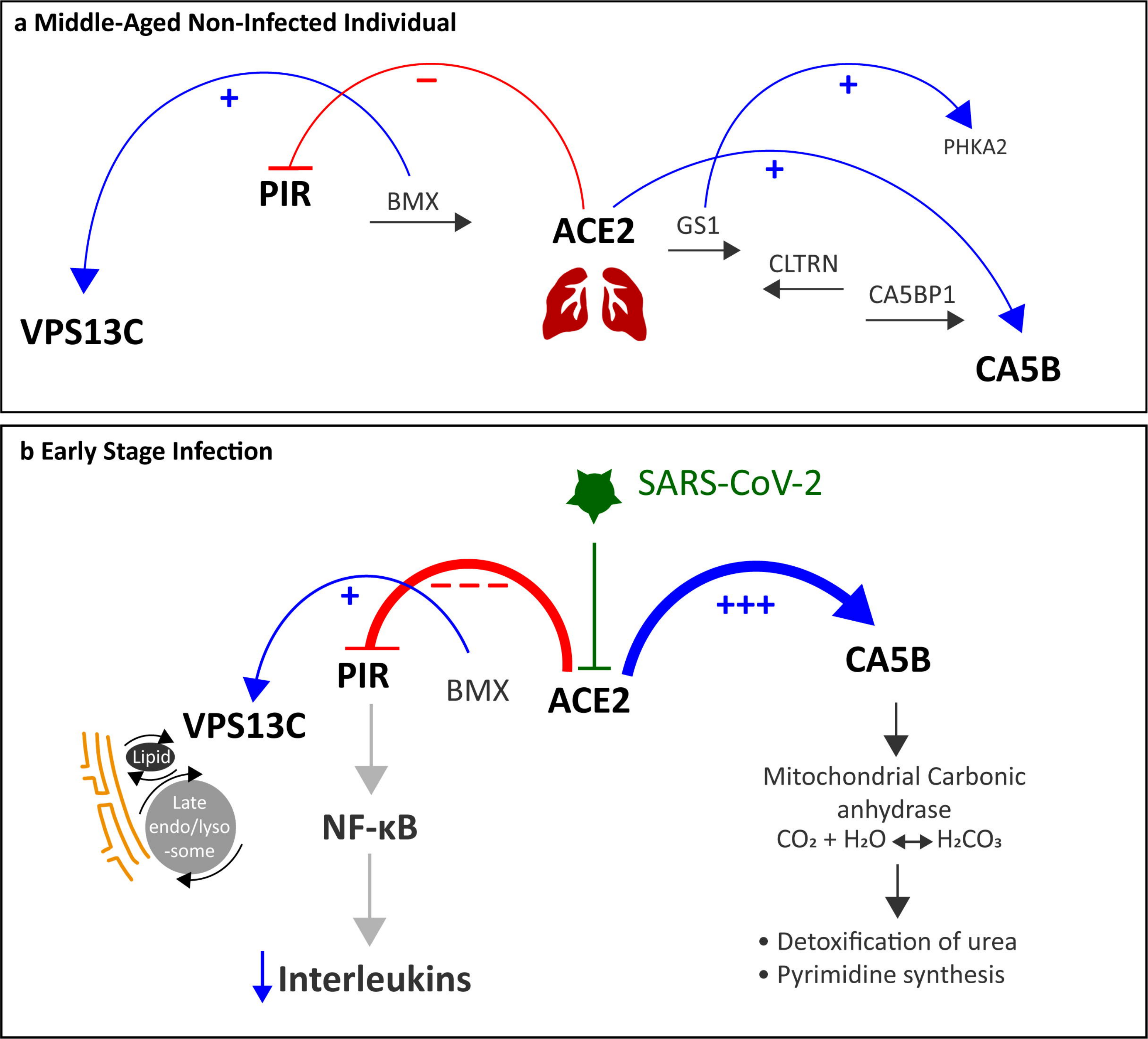
S4 Table. The common variants overlap DNA binding motifs. Data from Haploreg v4.1 (3/3/2020)

Supplementary Figures

S1 Figure. The CoDeS3D algorithm used in this study. Restriction fragments containing SNPs located within the *ACE* locus (chrX:15,519,996-15,643,106) were identified. Lung-specific Hi-C libraries were interrogated to identify genes in fragments that spatially interact (in cis- and trans-) with SNP-containing fragments. The identified spatial SNP-gene pairs were further used to query GTEx lung tissue (dbGaP Accession phs000424.v8.p2, UBERON:0008952). The Benjamini-Hochberg FDR control algorithm was applied to adjust the p values of the resulting eQTL associations to identify only significant (FDR < 0.05) lung-specific SNP-gene spatial relationships in the *ACE* locus.

S2 Figure. The eQTL data used in this study was obtained from lung samples taken from middle-aged individuals. To assess the correlation of genetic variation with the changes in gene expression, the GTEx project (https://gtexportal.org/home/) collected and analysed lung samples from donors who were densely genotyped. The age-distribution graph illustrates that approximately 70% of the lung samples that were obtained were from donors aged between 50 and 60.


A
IMR90_Rao2014
259
9
chrX
15,400,000
15,500,000
15,600,000
15,700,000
15,800,000
15,900,000
Scale
100 kb
hg38
Lung chromatin interactions
VPS13C
(chr15:61852389-62060473)
PHKA2
(chrX:18892300-18984598)
PIR
ACE2
CA5BP1
CA5B
AP1S2
GS1-594A7.3
BMX
CLTRN
ZRSR2
B
1.00
0.75
0.50
0.25
MAF
Population
AFR
AMR
ASN
EUR
chrX:
15560000
15570000
15580000
15590000
15600000
Scale
10 kb
hg38
SNPs_in_ACE2_eQTLs_of_CA5B
CA5B_eQTLs
SNPs_in_ACE2_eQTLs_of_PIR
PIR_eQTLs
100
Layered H3K27Ac
0
H3K27Ac Mark (Often Found Near Regulatory Elements) on 7 cell lines from ENCODE
DNase Clusters
DNase I Hypersensitivity Peak Clusters from ENCODE (95 cell types)
NCBI RefSeq genes, curated subset (NM_*, NR_*, NP_* or YP_*) - Annotation Release NCBI Homo sapiens Updated Annotation Release 109.20190905 (2019-09-10)
ACE2
ACE2


a Middle-Aged Non-Infected Individual
+
−
+
PHKA2
+
PIR
BMX
ACE2
GS1
CLTRN
CA5BP1
VPS13C
CA5B
b Early Stage Infection
SARS-CoV-2
+
− − −
+++
PIR
BMX
ACE2
CA5B
VPS13C
Lipid
Late endo/lyso-some
NF-κB
↓ Interleukins
Mitochondrial Carbonic anhydrase
$CO_2 + H_2O \leftrightarrow H_2CO_3$
• Detoxification of urea
• Pyrimidine synthesis